What Constitutes Happiness? Predicting and Characterizing the Ingredients of Happiness Using Emotion Intensity Analysis

Raj Kumar Gupta, Prasanta Bhattacharya, and Yinping Yang ✉

Institute of High Performance Computing (IHPC), Agency for Science, Technology and Research (A*STAR), Singapore, #16-16, 1 Fusionopolis Way, Singapore 138632
{gupta-rk,prasanta_bhattacharya,yangyp}@ihpc.a-star.edu.sg

Abstract. This paper explores the use of emotion intensity analysis in predicting and understanding the ingredients of happiness as expressed in text. We show that by using just the five dimensions of emotion intensity features (i.e., joy, anger, fear, sadness and overall valence), we can achieve good accuracies in classifying agency (i.e., whether or not the author of the happy moment is in control) (ACC=73.8%, AUC=.579, F1=.849) and in classifying social (i.e., whether or not the happy moment involves other people) (ACC=60.3%, AUC=.637, F1=.603). By integrating emotion intensity with sentiment, linguistics, demographics, concepts, and word embedding features, our final hybrid model performed significantly better for agency (ACC=83.5%, AUC=.887, F1=.893) and for social (ACC=90.3%, AUC=.959, F1=.907) predictions. Furthermore, we uncovered interesting patterns in how emotion intensities characterized happiness expressions across the various concepts (e.g., family, food, career, animals), between the two reflection periods (24 hours vs. 3 months), and across seven user-generated content corpora sources (HappyDB vs. MySpace, Runners World, Twitter, Digg, BBC and YouTube).

Keywords: emotion intensity, emotion analysis, affect, happiness, concept categories, reflection period, positive psychology

1 Introduction

While negative human emotions such as anger, fear or sadness are key subjects of study in the social sciences and humanities, the science of happiness has received growing attention in recent decades. Positive psychology is in itself a notable field of study focusing on factors that affect and improve happiness [3].

To provide a linguistic source for happiness research, Asai et al. [1] has developed a corpus of 100,000 happy moments—popularly known as HappyDB—based on crowdsourced contributions by Amazon Mechanical Turkers. In the CL-Aff shared task [9], a new corpus with 10,560 short text has been constructed as training data with annotated labels that identify the 'agency' of the author and the 'social' characteristic of the moment, in addition to demographic labels (e.g., age, gender, marital status) and concept labels describing its theme (e.g., romance, career, family, food).

The CL-Aff task data [1, 9] offers a rich linguistic source for the study of ingredients of happy moments. Of 10,506 happy moments, 73.8% (7,796) are labelled with 'yes' for agency (indicating that the author is in control), and 53.3% (5,625) are labeled with 'yes' for social (indicating that individuals other than the author are involved). Majority of the authors are single (53.4%), followed by married (41.9%), then divorced, separated and widowed (4.5%). 41.9% are female, and 38.8% have children. Majority of authors live in the U.S. (79.8%).

Though "happy moments" are inevitably positive feelings, do they vary in their degree or intensity of happiness? Our interest in this study focuses on the use of emotion intensity analysis to distinguish the happy moments. Furthermore, we study how this emotion intensity analysis contributes to the classification of the different characteristics—agency and social—of the happy moments.

2 Feature Extraction

Emotion Intensity (EI) Features. While humans experience and express emotions every day, the intensity of our emotions varies to a great extent [4, 13]. For example, within the course of a single event, we can experience a range of negative emotions from feeling slightly annoyed, to feeling extremely angry, or even raged. We may also feel worried or anxious, and sometimes even petrified. Positive emotions can also similarly range from feeling content, serene, proud and grateful, to higher-intensity joy, excitement and ecstasy. In the present context, for example, the expression "*Today is my last day of work before heading off for vacation, I am very happy and excited right now!*" expresses a greater amount of joy than "*I taught my neighbor how to change a flat tire*.", even though both are essentially instances of a joyous event.

We extracted the emotion intensity information for each line of happy moments using a tool we developed in a recent work [5]. The tool, named as *CrystalFeel*, is a collection of five SVM-based systems using features derived from parts-of-speech, n-grams, word embedding, and multiple affective lexicons, trained and validated using labelled tweets from the SemEval-18 affect in tweets task [10]. For each text, *CrystalFeel* automatically quantifies the intensity of five dimensions of emotions, i.e., valence, joy, anger, fear, and sadness, as real-valued scores. Evaluated on SemEval-18 test data, the tool has been shown to achieve Pearson correlations of .816, .708, .740, .700, .720 with ground truth data, for valence, anger, sadness, fear and joy respectively [5].

To illustrate, we excerpted five happy moments' examples from applying the tool on the CL-Aff training dataset (see Table 1). The first example is the one with highest valence intensity, and the remaining four are the ones with highest joy, anger, fear and sadness intensities respectively. It is useful to note that as emotion intensity is independently trained and developed for each of the five dimensions (i.e., valence, joy, anger, sadness, fear) with separate models, the comparison *within* each dimension is meaningful, but comparison *across* the dimensions is not meaningful or interpretable.

Table 1. Sample happy moments with highest *CrystalFeel* emotion intensity-scores

Happy Moments \ Emotion Intensity Scores	Emotion Intensity				
	valence	**joy**	**anger**	**fear**	**sadness**
LAST SUMMER HOLIDAYS WAS AN UNFORGETTABLE EXPERIENCE.MY FRIEND ALWAYS ENJOYED MANY PLACES.WE ENJOYED OUR EXCURSION WITH ENJOY AND BEAUTIFUL.I AM REMEMBER THAT SO EXCITING IN MY LIFE.	**1.181**[1]	.671	.314	.299	.267
Today is my last day of work before heading off for vacation, I am very happy and excited right now!	.811	**.729**	.377	.360	.417
Venting on the phone to the rep at Amazon for an order I placed and received the wrong item even though it didn't rectify the mistake	.320	.291	**.681**	.514	.515
I purchased a book for my son that I felt would help him with anxiety.	.366	.388	.499	**.668**	.534
I quit my job after feeling very depressed and I now feel relieved.	.365	.370	.523	.625	**.743**

Happy Moments with High- vs. Low-Emotion Intensities. The continuous emotion intensity scores can be converted to categorical values. For instance, using emotion intensity of 0.5 as a general threshold [10], an anger intensity of value > 0.5 can be considered to be "high-intensity" anger.

The following examples (first two with valence intensity > 0.5; second two <= 0.5) illustrate that the categorization of emotion intensity scores can provide an avenue to further distinguish happy moments, where the high-intensity ones tended to be associated with joyful emotions such as cheerfulness and excitement, whereas low-intensity ones were more linked to contentment and relief.

1) My favorite basketball team had a semi-critical game, and came out with a big win in exciting and entertaining fashion. (valence intensity = .849)
2) Last night, we had a reunion dinner with a group of good friends; the food, drinks, and company were amazing! (valence intensity = .846)
3) I didn't get fired from my job after a major screw up I caused, I managed to side step around the blame. (valence intensity = .306)
4) My dad getting out of the hospital after being sick for two weeks with pneumonia and an ear infection. (valence intensity = .282)

Sentiment, Personal Pronouns and other Linguistics (SPL) Features. We extracted a set of textual features using the R textfeatures package[2] including a combination of textual sentiment based on popular dictionaries by Nielsen [11] and Liu et al. [7], count of personal pronouns (e.g. count of first and second person mentions), and

---

[1] The score can be sometimes greater than 1 which indicates extremely high valence intensity is expressed in this text (higher than anyone in the system's training tweets data trained using linear regressions).

[2] https://textfeatures.mikewk.com

other linguistic features (e.g. count of words). These features were subsequently log-transformed and mean-centered before including them in our correlational analyses and classification models. There are a total of 24 dimensions in this feature set.

Demographics and Concepts (DC) Features. We also leveraged the demographics and concepts metadata provided in the training dataset. The data, originally provided as categorical values (e.g., concepts, parenthood, gender), was transformed into binary features. This led to a 29 dimensional feature set.

Correlation Analysis. To assess the effect of the individual features, we performed bivariate correlations analyses between the outcome variables and key features extracted for the model. Kendall's τ rank coefficients were used as the variables distributions do not conform to the normal distribution, and include binary variables.

The analysis indicates that emotion intensity scores are significantly associated with both agency and social labels. Agency is linked to higher *valence intensity* (τ=.016*), lower *joy intensity* (τ=-.025**), lower *anger* (τ=-.052**) and lower *fear* (τ=-.041**) intensities. Social is associated with higher *valence* (τ=.138**), higher *joy* (τ=.146**), lower *anger* (τ=-.086**) and lower *sadness* intensities (τ=-.102**).

For social, 12 features including emotion *valence* intensity (τ=.138**), *joy* intensity (τ=.146**), *sadness* intensity (τ=-.102**), *first person mentions* (τ=.240**), *number of characters* (τ=.166**), *words* (τ=.158**), *characters per word* (τ=.103**), *upper cases* (τ=-.192**), and *lower cases* (τ=.152*), *prepositions words* (τ=.131**), were found to be the most notable individual predictors.

Most of the other features in our model were also found to be statistically significant. Many of the correlations, however, were small in size. Appendix A provides detailed correlation results.

3 What are the Ingredients of Happiness?

Feature Experiments with 10-fold Cross Validation. For assessing the performance of the feature model, we first ran the three main feature sets separately to identify the individual performance, followed by a combination of these feature sets to explore the hybrid model's performance. All classifiers were based on logistic regression.

To complement the main features, we used FastText [6] and GloVe [12] embedding techniques to extract 200-dimensional word embedding features. To train the FastText features, we downloaded approximately 8 million tweets using Twitter Streaming API. For GloVe, we downloaded the model provided by the authors [12][3]. Table 3 and Table 4 provide the detailed experiment results.

With classifying agency, the overall pattern showed that the hybrid model led to better overall performances across *F1*, *AUC* and *ACC* measures. Notably, although the emotion intensity feature set has only five dimensions, it led to a comparable *F1* score (.849) as compared to the 100-dimensional FastText features (.884) and 100-dimensinoal GloVe (.879). The final hybrid model (258D) achieved the highest performance (*F1*=.893, *AUC*=.887, *ACC*=83.5%).

[3] http://nlp.stanford.edu/data/glove.6B.zip

Table 3. Feature experiment results for agency classification

Feature Sets	**Precision**	**Recall**	**F1**	**AUC**	**ACC**
Individual Feature Sets					
EI (5D)	.738	**1.000**	**.849**	.579	73.8%
SPL (24D)	.741	.992	**.849**	.636	73.9%
DC (29D)	**.750**	.976	.848	**.710**	**74.2%**
Hybrid Feature Sets					
EI + SPL (29D)	.743	**.987**	.848	.659	73.8%
EI + DC (34D)	.754	.971	.849	.715	74.5%
SPL + DC (53D)	.768	.961	.854	.732	75.7%
EI + SPL + DC (58D)	**.772**	.959	**.856**	**.739**	**76.1%**
Hybrid Feature Sets + Word Embedding (FastText and GloVe)					
FastText (100D)	.839	.934	.884	.872	81.9%
EI + FastText (105D)	.839	.934	.884	.872	81.9%
SPL + FastText (124D)	.843	**.936**	**.887**	.875	**82.5%**
DC + FastText (129D)	.845	.932	.886	.875	82.3%
EI + DC + FastText (134D)	.844	.933	.886	.875	82.3%
EI + SPL + DC + FastText (158D)	**.846**	.933	**.887**	**.878**	**82.5%**
GloVe (100D)	.832	.932	.879	.858	81.1%
EI + GloVe (105D)	.830	.931	.878	.858	80.1%
SPL + GloVe (124D)	.833	**.933**	.880	.861	81.3%
DC + GloVe (129D)	.836	.932	**.881**	.863	**81.5%**
EI + DC + GloVe (134D)	.836	.931	**.881**	.863	81.4%
EI + SPL + DC + GloVe (158D)	**.837**	.931	**.881**	**.865**	**81.5%**
Full Hybrid Model					
EI + SPL + DC + FastText + Glove (258D)	**.854**	**.936**	**.893**	**.887**	**83.5%**

Table 4. Feature experiment results for social classification

Feature Sets	**Precision**	**Recall**	**F1**	**AUC**	**ACC**
Individual Feature Sets					
EI (5D)	.642	.569	.603	.637	60.3%
SPL (24D)	.703	**.667**	.684	.740	67.3%
DC (29D)	**.950**	.639	**.764**	**.847**	**79.0%**
Hybrid Feature Sets					
EI + SPL (29D)	.717	.673	.694	.756	68.5%
EI + DC (34D)	**.927**	.665	.774	.865	79.4%
SPL + DC (53D)	.915	.711	.800	.887	81.1%
EI + SPL + DC (58D)	.914	**.721**	**.806**	**.890**	**81.6%**
Hybrid Feature Sets + Word Embedding (FastText and GloVe)					
FastText (100D)	.897	.849	.873	.933	86.8%
EI + FastText (105D)	.899	.855	.876	.935	87.2%
SPL + FastText (124D)	.904	.862	.882	.941	87.8%
DC + FastText (129D)	.922	.857	.888	.947	88.5%
EI + DC + FastText (134D)	.924	.859	.890	.948	88.7%
EI + SPL + DC + FastText (158D)	**.928**	**.870**	**.898**	**.952**	**89.5%**
GloVe (100D)	.902	.866	.884	.941	87.8%
EI + GloVe (105D)	.903	.869	.885	.943	88.0%
SPL + GloVe (124D)	.906	.873	.889	.946	88.4%
DC + GloVe (129D)	.922	.871	.896	.952	89.2%
EI + DC + GloVe (134D)	.922	.873	.897	.953	89.3%
EI + SPL + DC + GloVe (158D)	**.926**	**.880**	**.902**	**.956**	**89.9%**
Full Hybrid Model					
EI + SPL + DC + FastText + Glove (258D)	**.931**	**.884**	**.907**	**.959**	**90.3%**

For classifying social, a similar pattern emerged for the hybrid features vs. individual feature sets. However, individually, high dimensional FastText (*F1*=.849) and Glove (*F1*=.866) features performed significantly better than the low-dimensional emotional intensity features (*F1*=.603). As with the agency models, the final hybrid model (258D) achieved the highest performance (*F1*=.907, *AUC*=.959, *ACC*=90.3%).

Individual Features Ranking. To find out the relative predictive value of the individual features in our model, we also performed an individual feature ranking analysis. Variable importance analyses are common techniques with tree-based classification models, such as a random forest. Such measures of importance are broadly categorized into two types – accuracy-based importance and Gini/purity based importance. For accuracy-based importance analyses, the values of a particular variable are shuffled and the resultant decrease in accuracy is recorded as a measure of the variable's relative contribution or importance in the model. The Gini-based importance analyses is perhaps more specific to tree-based models as when a tree is built, the decision on which variable to branch on, at a particular node, is often based on a calculation of the Gini impurity or information gain metric.

Fig. 1 and Fig. 2 describe the impurity-based variable importance scores for the EI features (5D) and SPL features (24D). The importance scores reported are the Mean Decrease Impurity (MDI) scores for the Gini impurity [8] based on weighted impurity decreases for all nodes wherever the focal variable was used, averaged over all trees in a random forest. While the absolute values of the mean decrease scores are not important, the relative rankings of the scores provide an illustration of the variable importance.

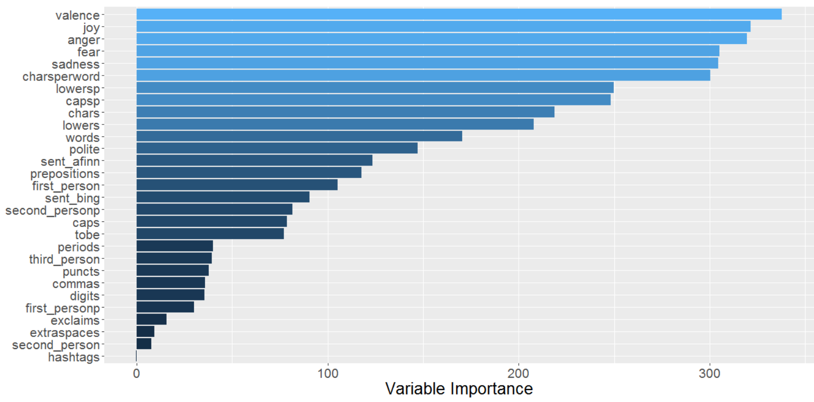

Fig. 1. Variable importance scores in classifying agency in moments

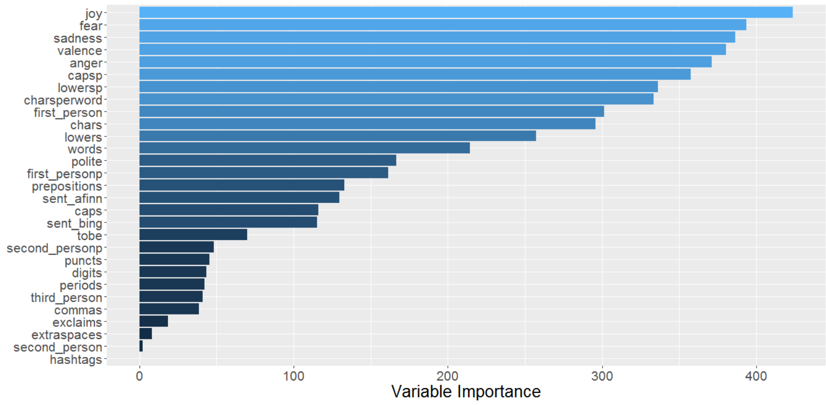

Fig. 2. Variable importance scores in classifying social in moments

The variable importance plots highlight that certain features consistently rank as most important features for both agency and social classifiers: the EI features, *number of characters*, *number of characters per word*, *proportion of lower case* and *upper case*. However, the relative importance of the emotion features varied for the classifiers. The *valence* score was found to be most predictive of agency, followed by intensity scores for *joy* and *anger*. In contrast, for the social classifier, *joy* was the most important predictor, followed by *fear* and *sadness*. Interestingly, some linguistic features (e.g., *first person mentions*) were more predictive of social than agency.

4 Understanding Happiness with Emotion Intensity

Emotion Intensity across HappyDB Concepts. To uncover further insights and characterize happy moments, focusing on emotion intensity, we analyzed the prevalence and intensity of various emotions across the 15 tagged concept categories. Fig. 3 shows the distribution of *valence* across the concepts, together with statistical significance results comparing the level of valence in each category.

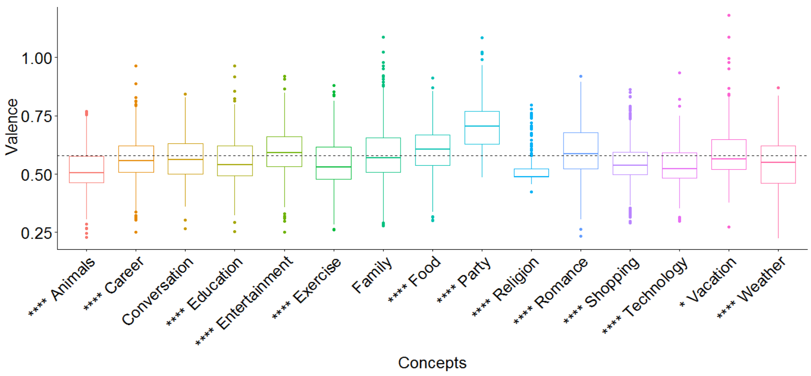

Fig. 3. Valence intensities across concepts (*p<.05, ** p<.01, *** p<.001, **** p<.0001)

We observe that the median *valence* scores for all concepts, except for *conversation* and *family,* are significantly different from the baseline (average) *valence*. While all categories within the HappyDB understandably exhibited positive *valence*, moments for the *party* category reported the highest *valence* relative to the baseline (t=26.052), followed by *food* (t=10.553), *entertainment* (t=7.187) and *romance* (t=5.920), while moments related to *animals* (t=-13.498), *shopping* (t=-10.979) and *religion* (t=-10.952) reported relatively lower *valence*.

Similarly, for the individual emotion categories of *joy*, *anger*, *fear* and *sadness*, we also observed a significant variance in intensities across the concepts. For instance, the intensity for *joy* was found to be high for *party* (t=16.324) and *family* (t=9.802), but relatively low for *career* (t=-11.654), *shopping* (t=-8.606), and *technology* (t=-8.396). Likewise, the intensity for *anger* was generally low across concepts, but particularly higher for *career* (t=16.391), and lower for *entertainment* (t=-16.799) and *party* (t=-21.569). For *fear*, we found significantly higher intensities for *religion* (t=15.361) and *exercise* (t=10.549), and lower intensities for *entertainment* (t=-14.604), *food* (t=-16.94) and *party* (t=-9.617). For *sadness*, the intensity was higher for *career* (t=20.015), and *religion* (t=9.579), but was substantially lower for *entertainment* (t=-25.265) and *party* (t=-16.401). All significances reported are smaller than .0001.

Emotion Intensity across Reflection Duration. We also analyzed the extent to which specific emotions were manifested in the happy moments for the two specific reflection durations, i.e., 24 hours vs. 3 months. No statistically significant difference was observed in the *valence* scores reported for the two reflection periods.

However, the analysis of specific emotions uncovered interesting differences. We found that the *anger* intensities were lower for the *3-month* than for the *24-hour* reflection period (t=3.091; p<.01) (Fig. 4). Similarly, the same pattern holds true for the *sadness* intensity (t=8.466; p<.001). No statistically significant difference was observed for the *joy* and *fear* emotions.

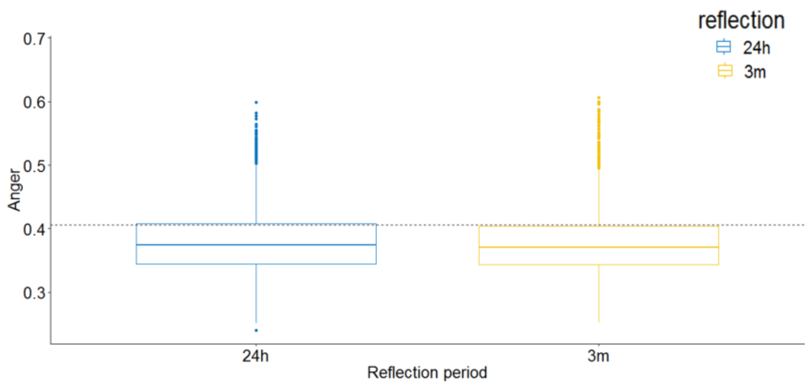

Fig. 4. Anger intensities across reflection periods

Plausibly, this pattern can be explained by the fact that when respondents were asked to recall happy moments from the past day, they were likely to cite fairly routine activities (e.g. watched a movie, had a dinner with family) where the presence of positive emotions might be mixed with certain negative ones as well (e.g. watched a tragic movie, but really enjoyed it). However, for longer time horizons, respondents were more likely to cite more significant events in their lives which are predominantly positive in their valence (e.g. birth of the first child, got promoted in job). In a future extension of this study, we plan to explore analyzing a collection of negatively valenced moments, i.e., an "unhappy" DB, for example, and reinvestigate the variance across reflection periods. We hypothesize that this difference between 3 months and 24 hours would be much more pronounced for negative emotional experiences. Specifically, the proportion of negative emotions for a 3 month reflection period would be much greater than for a 24 hour reflection period. This is because, as studies in psychology suggest, we tend to have stronger episodic memory of negative events, largely because these memories serve a certain evolutionary purpose [2]. The interplay between emotion categories and their recall characteristics is an interesting and important area of work that we plan to explore.

Emotion Intensity across Corpora. To compare the emotion intensity other corpora, we downloaded six datasets[5] [14] for a further analysis. Table 5 provides the mean statistics of EI scores for each of the corpora. Fig. 5 visualizes the EI scores using radar charts to profile the different corpora.

Table 5. Results of mean emotion intensity scores across corpora

	Mean words	Data size	valence	joy	anger	fear	sadness
CL-Aff (HappyDB)	13.38	**10,560**	.575	.490	.379	.364	.408
MySpace	2.08	1,041	**.593**	**.507**	.382	.338	.376
Runners World	**65.13**	1,046	.558	.455	.440	.426	.420
YouTube	17.12	3,407	.544	.460	.400	.366	.398
Twitter	15.35	4,218	.528	.449	.401	.369	.394
Digg	31.49	1,077	.465	.389	.449	.416	.432
BBC	62.54	1,000	.457	.360	**.474**	**.474**	**.448**

The emotion intensity profile analysis revealed interesting patterns. *HappyDB*'s profile showed an inclination towards high-intensity of valence and moderate-intensity of joy, which was most similar to the profile of *MySpace*, followed by *RW*, *YouTube* and *Twitter*. The emotion intensity profiles for *Digg* and *BBC* were found to be very different, in that they exhibited lower intensities of positive emotions and higher intensities of negative emotions. *BBC* was almost the complete opposite, showing a significant inclination towards highest anger, fear and sadness among corpora.

5 http://sentistrength.wlv.ac.uk/documentation/6humanCodedDataSets.zip

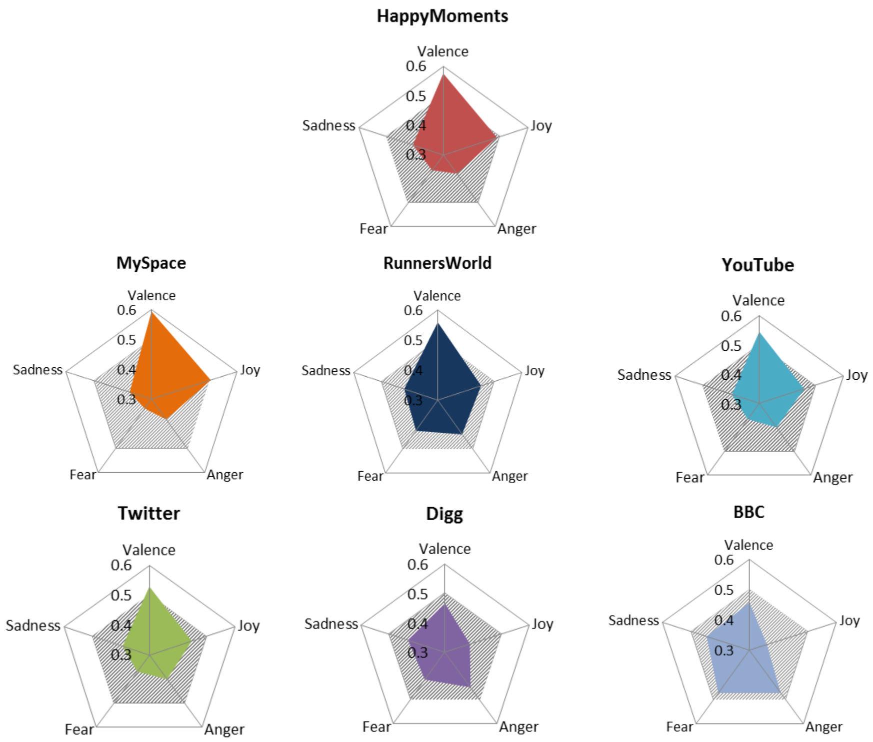

Fig. 5. Mean emotion intensity scores across six other corpora

5 Discussion and Conclusion

Emotion intensity is one of the most complex psychological constructs [4, 13]. Our main contribution to this challenge is to introduce the notion of emotion intensity as a useful factor to account for, in happiness research. We found that happiness expressions can be effectively distinguished by valence, joy, anger, fear, and sadness intensity scores [5]. We described our system incorporating emotion intensity as a main, psychologically meaningful feature set for classifying agency and social, and reported feature experiments showing how they contribute to the system's performance.

For the open task, our analyses overlaying emotion intensity with the various other happiness dimensions revealed interesting new patterns. For instance, moments associated with *party* were higher in *joy*, but significantly lower in all other emotions. We also found that the intensities of anger and sadness were significantly lower across the *3-month* than the *24-hour* reflection period. Lastly, profiling various corpora revealed interesting differences: *HappyDB* is closer to *MySpace*, *Runners World, YouTube* and *Twitter*, but is quite different from comments under media sites *Digg* and *BBC*.

The emotion intensity scores appear to characterize happiness effectively, especially in conjunction with other dimensions such as concepts and reflection period. In future work, we plan to explore how the emotion intensity measures interact with other dimensions such as demographics. It might be also interesting to broaden the scope of happiness research by cross-studying how happy moments differ from sad, angry and fearful moments, through which the community can potentially uncover richer insights underlying important human experiences.

Acknowledgement

This research is supported by SERC Strategic Fund from Science and Engineering Research Council (SERC), A*STAR (project no. a1718g0046).

References

1. Asai, A., Evensen, S., Golshan, B., Halevy, A., Li, V., Lopatenko, A., Stepanov, D., Suhara, Y., Tan, W-C, Xu, Y.: HappyDB: A Corpus of 100,000 Crowdsourced Happy Moments, arXiv:1801.07746v2 [cs.CL] (2018)
2. Baumeister, R. F., Bratslavsky, E., Finkenauer, C., Vohs, K. D.: Bad is stronger than good. Review of General Psychology, 5(4), 323. (2001)
3. Compton, W.C., Hoffman, E.: Positive Psychology: The Science of Happiness and Flourishing. 2nd ed. Belmont, CA: Wadsworth Cengage Learning, (2013)
4. Frijda, N.H., Ortony, A., Sonnemans, J., Clore, G.: The complexity of intensity: Issues concerning the structure of emotion intensity. Personality and Social Psychology Review 13:60 – 89. (1992)
5. Gupta, R.K. and Yang, Y.: CrystalFeel at SemEval-2018 Task 1: Understanding and Detecting Intensity of Emotions using Affective Lexicons, 12th International Workshop on Semantic Evaluation (SemEval 2018). (2018)
6. Joulin, A. Grave, E., Bojanowski, P., Mikolov, T.: Bag of Tricks for Efficient Text Classification. European Chapter of the Association for Computational Linguistics, 2, pp. 427-431 (2016)
7. Liu, B., Hu, M., Cheng, J.: Opinion observer: analyzing and comparing opinions on the web. 14th international conference on World Wide Web, (2005)
8. Louppe, G., Wehenkel, L., Sutera, A., Geurts, P.: Understanding variable importances in forests of randomized trees. In Advances in neural information processing systems. pp. 431-439 (2013)
9. Jaidka, K., Mumick, S., Chhaya, N., Ungar, L.: The CL-Aff Happiness Shared Task: Results and Key Insights, Proceedings of the 2nd Workshop on Affective Content Analysis @ AAAI (AffCon2019). (2019)
10. Mohammad, S.M., Bravo-Marquez, F., Salameh, M., Kiritchenko, S.: Semeval-2018 task 1: Affect in tweets. SemEval. (2018)
11. Nielsen, F.Å.: A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. arXiv preprint arXiv:1103.2903 (2011)
12. Pennington, J., Socher, R., Manning, C.D.: GloVe: Global Vectors for Word Representation. Empirical Methods in Natural Language Processing (EMNLP), pp. 1532 – 1543 (2014)

13. Sonnemans, J., Frijda, N. H.: The structure of subjective emotional intensity. Cognition and Emotion, 8(4), pp. 329-350. (1994)
14. Thelwall, M., Buckley, K., Paltoglou, G.: Sentiment strength detection for the social Web, Journal of the American Society for Information Science and Technology, 63(1), pp. 163-173. (2012)

Appendix A: Bivariate correlations analysis results

Features group	**Features code**	**Description**	**Correlation coefficients (Kendau's τ)**	
			agency	**social**
Emotion Intensity (5D)	valence	Intensity of overall valence	.016*	**.138****
	joy	Intensity of joy	-.025**	**.146****
	anger	Intensity of anger	-.052**	-.086**
	fear	Intensity of fear	-.041**	.006
	sadness	Intensity of sadness	-.014	**-.102****
Sentiment (2D)	sent_afinn	AFINN sentiment score	-.041**	.078**
	sent_bing	Bing sentiment score	-.036**	.039**
Personal Pronouns (5D)	first_person	# of first person mentions	-.034**	**.240****
	first_personp	% of first person mentions	-.025*	**.211****
	second_person	# of second person mentions	-.033**	.036**
	second_personp	% of second person mentions	**-.138****	.095**
	third_person	# of third person mentions	-.057**	.025*
Other Linguistics (17D)	chars	# of characters	-.094**	**.166****
	words	# of words	-.079**	**.158****
	charsperword	# of characters per word	-.095**	**.103****
	caps	# of upper case	.041**	-.092**
	capsp	% of upper case	**.105****	**-.192****
	lowers	# of lower case	-.087**	**.152****
	lowersp	% of lower case	-.075**	**.166****
	puncts	# of punctuation marks	-.046**	.057**
	periods	# of periods	-.020*	.056**
	commas	# of commas	-.034**	.031**
	exclaims	# of exclamation marks	-.028**	-.028**
	prepositions	# of prepositions	.009	**.131****
	digits	# of numeric mentions	-.025**	-.024*
	extraspaces	# of extra/unrequired spaces	-.012	.025*
	polite	# of politeness-related words	-.016	.068**
	tobe	# of mentions of 'to-be'	-.092**	.030**
	hashtags	# of hashtags mentioned	-.003	.005

* Correlation is significant at the 0.05 level (2-tailed); ** at the 0.01 level (2-tailed)